(12) UK Patent Application (19) GB (11) 2 284 911 (13) A

(43) Date of A Publication 21.06.1995

(21) Application No 9325747.5

(22) Date of Filing 16.12.1993

(71) Applicant(s)
Plessey Semiconductors Limited

(Incorporated in the United Kingdom)

Cheney Manor, SWINDON, Wiltshire, SN2 2QW, United Kingdom

(72) Inventor(s)
Neil Stuart Hastie

(74) Agent and/or Address for Service
Maurice Edward Casey
GEC Patent Department, Waterhouse Lane, CHELMSFORD, Essex, CM1 2QX, United Kingdom

(51) INT CL[6]
G06F 12/08

(52) UK CL (Edition N)
G4A AMC

(56) Documents Cited
EP 0325420 A2

(58) Field of Search
UK CL (Edition M) G4A AMC
INT CL[5] G06F 12/08
ONLINE DATABASE : WPI

(54) Flexible lock-down cache.

(57) A flexible lock-down cache memory arrangement in which any required number of storage locations in a cache memory may be marked as "locked-down" for use as static random access memory for critical program elements, such as interrupt routines, which require the fastest possible access.

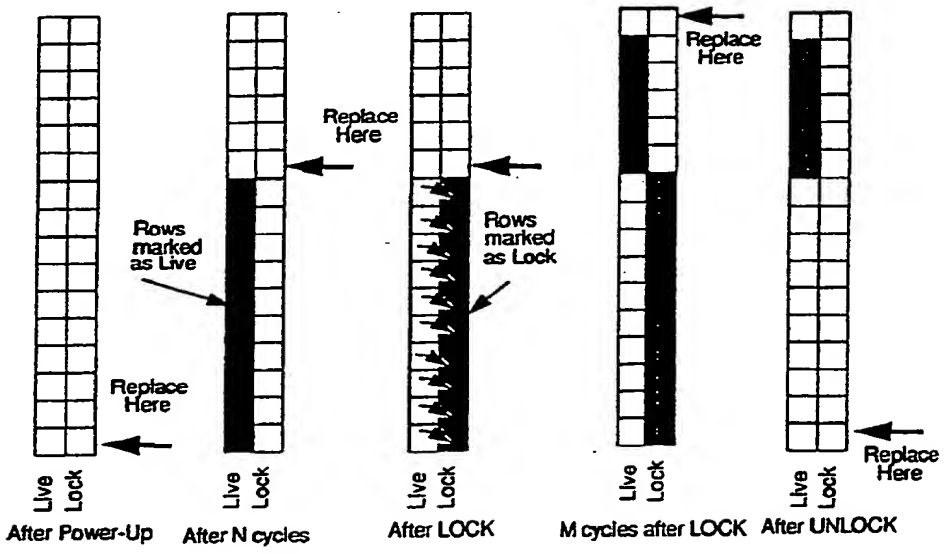
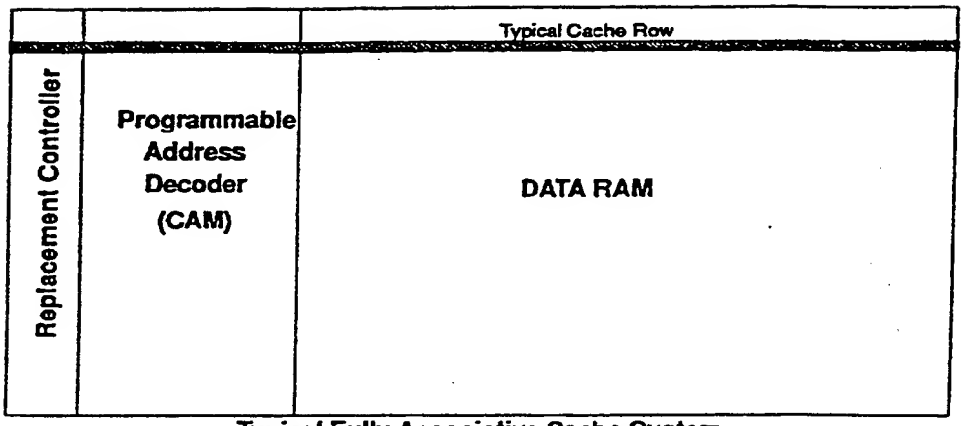


Fig 2

GB 2 284 911 A



Typical Fully Associative Cache System

Fig 1

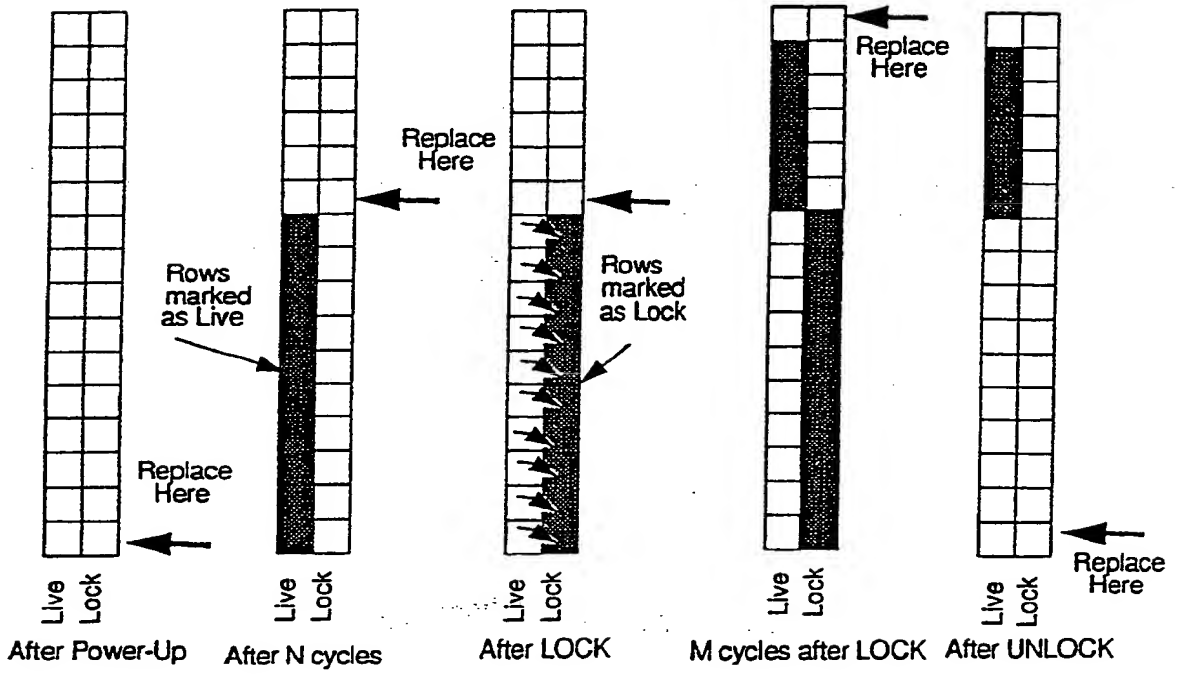


Fig 2

DATA STORAGE ARRANGEMENT

Most commercial processors now have cache memory systems integrated on chip. These are memory systems that provide quick access to the most frequently or recently used code and/or data items and hence speed programme execution. In many real time systems however it is not the most recently used code segments that require the fastest access time, but the most critical code elements such as interrupt routines etc. To store these critical code elements an on chip SRAM may be used, into which the critical code elements are loaded prior to use.

Several cached processor designs provide the ability to use a portion of the cache as SRAM. These caches are known as a "lock down" caches. The split within the cache between standard cache and the lock down SRAM area is invariably on a large boundary. Thus typically a 8KByte memory may be arranged as a 4K cache plus 4K SRAM or a 6K cache plus 2K SRAM. The drawback with this approach is that the size of memory can only be coarsely split between the two memory types. For instance in the above example if the critical code requires 2.1KBytes then the available memory must be split as a 4K/4K to accommodate it. This has the effect of limiting the amount of cache available whilst providing far too much SRAM.

It is an object of the present invention to provide a flexible lock down cache system allowing the user to precisely allocate the Cache/SRAM split according to the current system requirements.

According to one aspect of the present invention in a data storage arrangement in which a data packet or word is arranged to be written in to an available one of a plurality of locations in a data store in association with an identifier by means of which said one of said plurality of locations may be selected when read-out of said data packet or word is required, and in which there are provided first means to indicate which of said plurality of locations is or are available, or become available at any time in accordance with a predetermined rule, for the writing or overwriting therein of a different data packet or word, there are provided second means selectively to mark one or more of said locations in said data store in which a data packet or word has been written as unavailable for overwriting with a different data packet or word.

According to another aspect of the present invention in a data storage arrangement comprising main memory and fast-access cache memory, in which a data packet or word is written into an available one of a plurality of locations in said cache memory in association with an identifier by means of which said one of said locations may be identified when read-out of said data packet or word is required, there are provided first means to select into which of the available ones of said plurality of locations in said cache memory a new data packet or word is to be written and second means selectively to mark one or more of said locations into which respective data packets or words have been written as not available for selection by said first means.

Said first means may be arranged to set one bit position of a location to a predetermined binary value when a data packet or word is written into that location and to reset that bit position or to repeat the setting in response to predetermined conditions of usage of said cache memory. The second means may be arranged to set a second bit position of a location to said predetermined binary value to mark that location as not available for selection by said first means.

A cache memory type of data storage arrangement in accordance with the present invention will now be described with reference to the drawing, in which:-

Figure 1 shows a cache memory system diagrammatically, and

Figure 2 illustrates the operation of such a memory system.

Referring to the drawing, the replacement algorithm proposed for such arrangements is a modification of the GODS algorithm. A "Live" bit in the first column of bit positions indicates the availability of the row or location for replacement. At the start of operation the cache starts filling from the bottom, and as each row is filled this bit is set indicating that the row is no longer available. Once the cache is completely filled these bits are cleared. A subsequent access to a row causes the "Live" bit to be set again as does a complete replacement of that row with a new address and data set. A pointer moves up the array indicating the next available position for replacement. To decide whether a row may be replaced the "Live" bit must be clear. As indicated in Figure 1 each cache row has a number of bits for data storage, a number of bits of programmable address and a number of bits to

control the replacement of the row within the cache.

To provide Lock down areas within the cache a second bit (the "Lock" bit) is added to the "Live" bit to indicate that the row is locked. A number of different algorithms are possible to set and reset this bit, a simple version is now described.

The critical data/instructions are loaded into the cache. This sets the "Live" bits for each row accessed. Following this a LOCK signal is exerted. This transfers the "Live" bits into a second column of bits, the "Lock" bits, arranged once again as one per row, as shown in Figure 2. The replacement algorithm must now interrogate both bits to decide whether a row may be replaced:- the live bit, and lock bit. If both bits are clear then the row may be replaced. To remove the locked area the "Lock" bits are merely reset and then replaced as described above.

Using this system any number of rows in the cache may be locked-down. This provides a variable size SRAM that may be modified as system demands change.

At a system level it would be very desirable to be able to pass to the cache a data item and tell it to lock it. Equally desirable would be the ability to unlock this particular item at a later point in time. (As opposed to unlocking all locked items in the cache). This can be achieved in the following manner. An address is passed to the cache controller, indicating that data found at this address should be loaded and locked. (The load is performed in the same way as a standard load in the case of a cache miss). The lock is indicated by a single bit on the cache row as described above. To unlock the data the processor passes the cache an address with instruction to unlock if present.

Other additional functions may be of use, such as passing the cache an address and the cache responding with information such as in cache/not in cache, locked/not locked etc. Other variations such as an ability to load and lock an address range, or to start and stop locking (to load a programme routine or a data structure) can easily be implemented.

This system has the following advantages over currently available commercial solutions:-

The amount of memory available to be locked is variable on a row by row basis, as

opposed to a more conventional block level.

Each row of data may be individually locked or unlocked.

The amount of memory available to be locked is dynamically variable as the programme execution proceeds, as opposed to being one time selectable at system start up.

The locked data may be from completely unrelated areas of the processor memory map due to the programmable (CAM) decoding structure used. In an SRAM based system the fixed decode scheme requires lock data to be accessed from sequential memory locations.

CLAIMS

1. A data storage arrangement in which a data packet or word is arranged to be written in to an available one of a plurality of locations in a data store in association with an identifier by means of which said one of said plurality of locations may be selected when read-out of said data packet or word is required, and in which there are provided first means to indicate which of said plurality of locations is or are available, or become available at any time in accordance with a predetermined rule, for the writing or overwriting therein of a different data packet or word, there are provided second means selectively to mark one or more of said locations in said data store in which a data packet or word has been written as unavailable for overwriting with a different data packet or word.

2. A data storage arrangement comprising main memory and fast-access cache memory, in which a data packet or word is written into an available one of a plurality of locations in said cache memory in association with an identifier by means of which said one of said locations may be identified when read-out of said data packet or word is required, there are provided first means to select into which of the available ones of said plurality of locations in said cache memory a new data packet or word is to be written and second means selectively to mark one or more of said locations into which respective data packets or words have been written as not available for selection by said first means.

3. A data storage arrangement in accordance with Claim 2 wherein said first means is arranged to set one bit position of a location to a predetermined binary value when a data packet or word is written into that location and to reset that bit position or to repeat the setting in response to predetermined conditions of usage of said cache memory.

4. A data storage arrangement in accordance with Claim 3 wherein said second means is arranged to set a second bit position of a location to said predetermined binary value to mark that location as not available for selection by said first means.

5. A data storage arrangement substantially as hereinbefore described with reference to the accompanying drawing.

Patents Act 1977 Examiner's report to the Comptroller under Section 17 (The Search report)	Application number GB 9325747.5
Relevant Technical Fields (i) UK Cl (Ed.M) G4A (AMC) (ii) Int Cl (Ed.5) G06F 12/08	Search Examiner S J PROBERT Date of completion of Search 4 FEBRUARY 1994
Databases (see below) (i) UK Patent Office collections of GB, EP, WO and US patent specifications. (ii) ONLINE DATABASE: WPI	Documents considered relevant following a search in respect of Claims :- 1-5

Categories of documents

X:	Document indicating lack of novelty or of inventive step.	P:	Document published on or after the declared priority date but before the filing date of the present application.
Y:	Document indicating lack of inventive step if combined with one or more other documents of the same category.	E:	Patent document published on or after, but with priority date earlier than, the filing date of the present application.
A:	Document indicating technological background and/or state of the art.	&:	Member of the same patent family; corresponding document.

Category	Identity of document and relevant passages		Relevant to claim(s)
X	EP 0325420 A2	(AMD) see column 3 line 63 to column 4 line 2; column 4 lines 35-42 and column 18 lines 27 to 40	1-5

Databases: The UK Patent Office database comprises classified collections of GB, EP, WO and US patent specifications as outlined periodically in the Official Journal (Patents). The on-line databases considered for search are also listed periodically in the Official Journal (Patents).